

Collation of Armenian manuscripts

A lone historian's approach

Գ, լե տալան և այն Դ-ից հետո Երթան գովեստից մարդկան։ Որոճալով որոճեն զգիւական եր
Հայաստանէսց. Ժ, ի վեր հայ

Tara L Andrews | University of Oxford | 18 November 2009

«**Ի**րերս, Կառնապի պատրենն զկարգ ժամատեղացն: լինին և բազումք ուսում

Ռեզնանի վարդապետանայ. ոչ արեղայ և ոչ կրանաւոր, ոչ իշխան և ոչ ժողովուրդ. Ի

Chronicle of Matthew of Edessa

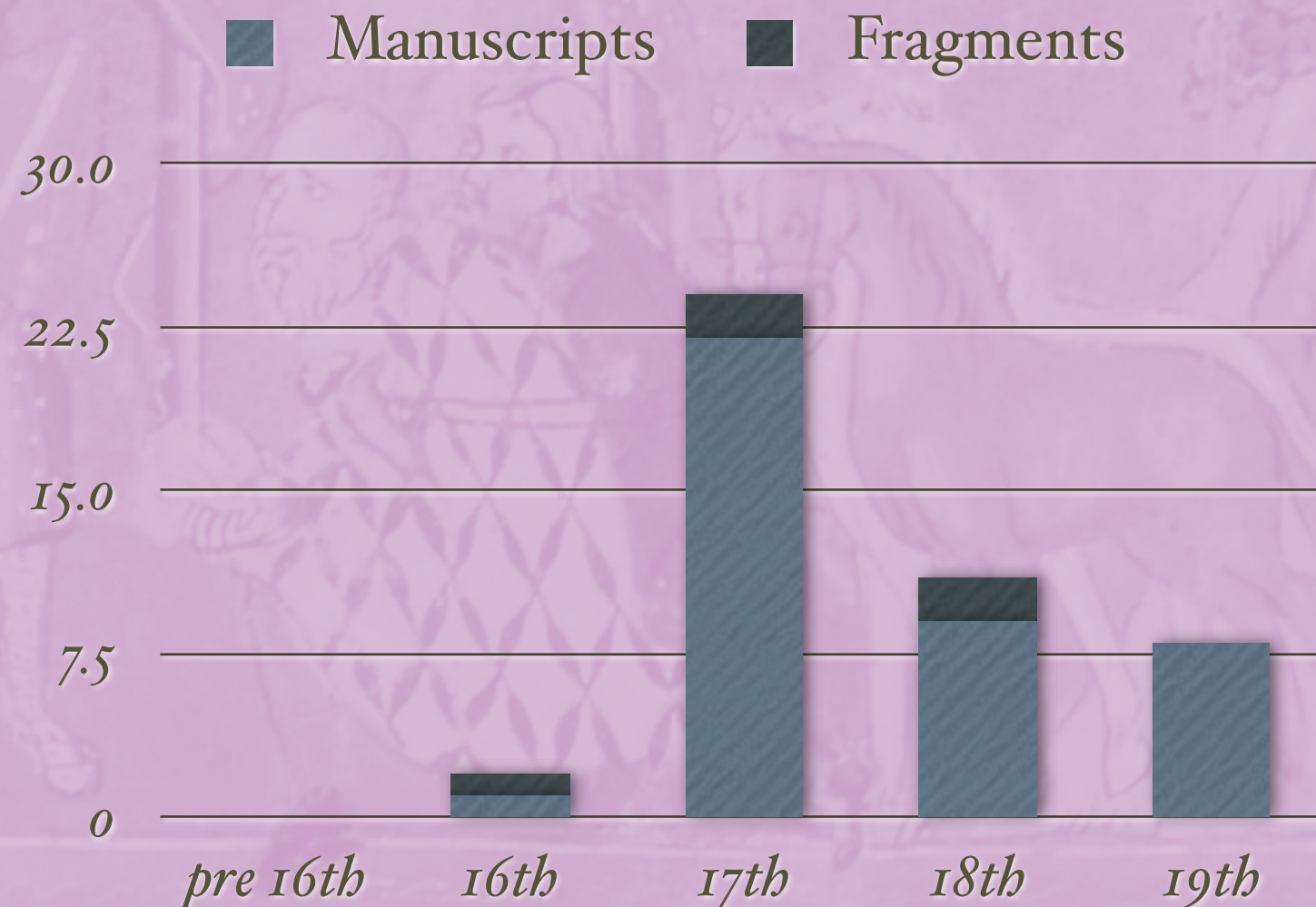
- Written around 1137
- 35 known manuscripts
- No manuscripts older than 1600

յ, Լե Կապուլան Եւրե Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե
Հայաստանէսց. Ժ, Ի վեր Կապուլան Եւրե Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե

Tara L Andrews | University of Oxford | 18 November 2009

Եւրե Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե
Հայաստանէսց. Ժ, Ի վեր Կապուլան Եւրե Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե

Extant manuscripts of the Chronicle



Գ, ԼԵ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ
ՀԱՅԱՍՏԱՆԻՆԵՅԵ. Ժ, ԻՎԵՐ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս» Երկրորդ, մեծագոյն լիարժէք զկարգ ժամատեղացն: լինին և բազումք ուսում

- No single manuscript can be the base
- No need to collate 'derivative' manuscripts
 - ...but difficult to identify these.
- Need solution: baseless collation and cladistic analysis.

Tara L Andrews | University of Oxford | 18 November 2009

- Way to collate transcribed manuscripts
- Way to review results for editorial decisions
 - Help keeping my decisions consistent
- Publication of finished edition

Specific collation requirements

- Unicode for Armenian script
- TEI files for manuscript transcriptions
- Output suitable for stemmatic analysis, direct publishing, or manipulation by other programs

յ, Ե ԿԱՎԱՆ ԵՆԵ ԴԱՎԱՆ ԵՐԺԱՆ ԳՈՎԵԱՄԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻԱԿԱՆ ԵՐԺ
ՀԱՅԱՍՏԱՆԻՆԵՅԵՅ. Ժ, Ի ՎԵՐ ՀԱՅ

Tara L Andrews | University of Oxford | 18 November 2009

ԻՆԱ "ԻՐԿՐՈՐԿ, ՎԱ ՄԱՍԿԱՄԱՍ ԳԱՊԻՄՈՐԵՆ ՂԿԱՐԳ ԺԱՄԱՄԵՂԱԿՆ: ԼԻՆԻՆ և ԲԱԳՈՒՄՔ ՈՒՍՈՒՄ

- COLLATE
- Juxta
- TUSTEP

Available collation software

- Juxta
- TUSTEP

Գ, ԼԵ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ
ՀԱՅԱՍՏԱՆԻԿԱՆ Ժ, ԻՎԵՐ ԿԱՆԿԱՆ

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս» Երկրորդ, մեծագոյն լիարժէք զկարգ ժամատեղացն: լինին և բազումք ուսում

Available collation software

- TUSTEP

Գ, ԼԵ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ
ՀԱՅԱՍՏԱՆԻԿԱՆ Ժ, ԻՎԵՐ ԿԱՆԿԱՆ

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս» Երկրորդ, մեծագոյն լիարժէք զկարգ ժամատեղացն: լինին և բազումք ուսում

Available collation software

Գ, ԼԵ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ
ՀԱՅԱՍՏԱՆԻԿԱՆ Ժ, ԻՎԵՐ ՀԱՅ

«Ինչպէս որ, Դ. Կարապետ Կարապետեան զկարգ ժամատեղացն: լինին և բազումք ուսում

Virtues of a programmer

- Laziness
- Impatience
- Hubris

- Unicode for Armenian script
- TEI files for manuscript transcriptions
- Output suitable for stemmatic analysis, direct publishing, or manipulation by other programs

- Unicode for Armenian script
- TEI files for manuscript transcriptions
- Output suitable for stemmatic analysis, direct publishing, or manipulation by other programs
- As little manual work as possible

Պէպէտէս Վարդապետեանայ. ոչ արեղայ և ոչ կրանաւոր, ոչ իշխան և ոչ ժողովուրդ. Ի

Re-inventing the wheel

- ‘Best-effort’ collation algorithm
- Based on diff
- Was spared the ‘transposition problem’

ց, Լէ Կապուան Եոյն Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական եր
Հայաստանէսց. Ժ, Ի վեր հաշ-

Tara L Andrews | University of Oxford | 18 November 2009

տեա "Իրհրորդ, Գա առաջադասութիւն պղտորեն զկարգ ժամատեղացն: լինին և բազումք ուսում

«**Ի**րերս, Կառնապի պատրենն զկարգ ժամատեղացն: լինին և բազումք ուսում

The white cat was sitting on top of a dining table.

SAME: The white cat was	ADD: + gingerly	SAME: sitting	CHANGE: - atop --- + on + top + of	SAME: a	DELETE: - dining	SAME: table
-------------------------------------	--------------------	------------------	---	------------	---------------------	----------------

Transposition

The quick brown fox jumped over the lazy dogs.
The lazy brown fox jumped over the quick dogs.

CHANGE: - quick --- + lazy	CHANGE: - lazy --- + quick
-------------------------------------	-------------------------------------

յ, Լէ Կարգաւոր Եւրոպայի Գեղեցիկ մարդկան: Որոճալով որոճեն զդիւական երե
հայտաւորանէ-յց. Ժ, Ի վեր- հաշ-

Ինչ "Իրհրորդ, Գա անարգաւոր պղտորեն զկարգ ժամատեղացն: Լինին և բազումք ուսում

Word movement

The quick brown fox jumped over the lazy dogs.
The quick fox jumped over the lazy brown dogs.

DELETE: - brown	ADD: + brown
--------------------	-----------------

յ, Լէ Կարգաւ Եւրէ Դարձեալ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե
Հայաստանէնց. Ժ, Ի վեր հաշ-

Tara L Andrews | University of Oxford | 18 November 2009

տեա "Իրհրորդ, Գա առաջաւոր Գա պղտորեն զկարգ ժամատեղացն: լինին և բազումք ուսում

In 401 of the Armenian era, there was a great famine in many places.

Գ, լե տալան և այն Դ-ից հետո Երթան գովեստից մարդկան։ Որոճալով որոճեն զգիւական եր
Հայաստանէսց. Ժ, ի վեր հայ

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս» Երկրորդ, մեծագոյն պատերէն զկարգ ժամատեղացն: լինին և բազումք ուսում

In 401 of the Armenian era, there was a great famine in many places.
In the Armenian year 401, there was a plague in many places.

Գ, լե տալան և այն հարկէր Երթան գովեստից մարդկան։ Որոճալով որոճեն զգիւական եր
հայտաւորուեց. Ժ, ի վեր հաշ-

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս» Երկրորդ, մեծագոյն պատերէն զկարգ ժամատեղացն: լինին և բազումք ուսում

Constructing the 'superbase'

In				401	of	the	Armenian	era,	there	was	a	great	famine	in	many	places.
In	the	Armenian	year	401,					there	was	a		plague	in	many	places.

Tara L Andrews | University of Oxford | 18 November 2009

«**Ի**րերս, Կառնապի պատրենն զկարգ ժամատեղացն: լինին և բազումք ուսում

Ռեզնանսի վարդապետանայ. ոչ արեղայ և ոչ կրանաւոր, ոչ իշխան և ոչ ժողովուրդ. Ի

Using the superbase

<i>In</i>	<i>the</i>	<i>Armenian</i>	<i>year</i> <i>era</i>	401	<i>of</i>	<i>the</i>	<i>Armenian</i>	<i>era,</i>	<i>there</i>	<i>was</i>	<i>once</i>	<i>more</i>	<i>a</i>	<i>great</i>	<i>famine</i> <i>plague</i>	<i>in</i>	<i>many</i>	<i>places.</i>
-----------	------------	-----------------	---------------------------	-----	-----------	------------	-----------------	-------------	--------------	------------	--------------------	--------------------	----------	--------------	--------------------------------	-----------	-------------	----------------

In				401	of	the	Armenian	era,	there	was			a	great	famine	in	many	places.
In	the	Armenian	year	401,					there	was			a		plague	in	many	places.
		Armenian	era	401,					there	was	once	more	a		plague.			

յ, Լե Կապուտան Երեւոյ Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երեւոյս
Հայաստանէսցոյ. Ժ, Ի վեր Կապուտան Երեւոյ Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երեւոյս

Tara L Andrews | University of Oxford | 18 November 2009

Երեւոյ Դաւիթ Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երեւոյս

Straight to TEI

```
<app>
  <rdg wit="#A #B"><w>In</w></rdg>
  <rdg wit="#C"/>
</app>
<app>
  <rdg wit="#B"><w>the</w></rdg>
  <rdg wit="#A #C"/>
</app>
<app>
  <rdg wit="#B #C"><w>Armenian</w></rdg>
  <rdg wit="#A"/>
</app>
<app>
  <rdg wit="#B"><w>year</w></rdg>
  <rdg wit="#C"><w>era</w></rdg>
  <rdg wit="#A"/>
</app>
<w>401</w>
<app>
  <rdg wit="#A"><w>of</w></rdg>
  <rdg wit="#B #C"/>
</app>
```

յ, Լե Կարդան Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե
հայապանէսց. Ժ, Ի վեր հաչ

Tara L Andrews | University of Oxford | 18 November 2009

Ինչպէս, Գա անպատիւ պղտորեն զկարգ ժամատեղացն: լինին և բազումք ուսում

It isn't perfect

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս» Երկրորդ, մեծագոյն լիարժէք զկարգ ժամատեղացն: լինին և բազումք ուսում

...but it was good enough.

Գ, ԼԵ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ
ՀԱՅԱՍՏԱՆԻՆԵՅԵՅ. Ժ, ԻՎԵՐ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ

«Ինչպէս որ, զանազան պատճառներով զկարգ ժամատեղացն: լինին և բազումք ուսում

- Use the collation results for stemmatic analysis
- Use the collation results to edit the text
- Send the collation results to another program for visualization

How not to do stemmatic analysis

Tara L Andrews | University of Oxford | 18 November 2009

Friday 28 February 14


```
accept (#)
accept detail (#)
subst detail (#) (value)
spelling (# alternate) (# canonical)
orth (# alternate) (# canonical)
emend (comment text)
note (note text)
next
help
save
quit
--> 
```


Գ, լե տալան և այն հարկէր Երթան գովեստից մարդկան։ Որոճալով որոճեն զդիւական եր
հայտաւորուեց. Ժ, ի վեր հաշ-

From collation results...

```
<app>
  <rdg wit="#A #B"><w>In</w></rdg>
  <rdg wit="#C"/>
</app>
<app>
  <rdg wit="#B"><w>the</w></rdg>
  <rdg wit="#A #C"/>
</app>
<app>
  <rdg wit="#B #C"><w>Armenian</w></rdg>
  <rdg wit="#A"/>
</app>
<app>
  <rdg wit="#B"><w>year</w></rdg>
  <rdg wit="#C"><w>era</w></rdg>
  <rdg wit="#A"/>
</app>
<w>401</w>
<app>
  <rdg wit="#A"><w>of</w></rdg>
  <rdg wit="#B #C"/>
</app>
```

յ, Լե Կարդան Երթան գովեստից մարդկան: Որոճալով որոճեն զդիւական եր
Հայաստանէնց. Ժ, Ի վեր հաշ

Tara L Andrews | University of Oxford | 18 November 2009

Ինչպէս, Գա անտառի մէջ պղտորեն զկարգ ժամատեղացն: Լինին և բազումք ուսում

Ռեզնանսի վարդապետանայ. ոչ արեղայ և ոչ կրանաւոր, ոչ իշխան և ոչ ժողովուրդ. Ի

...to edited text

```
<app>
  <lem wit="#A #B"><w>In</w><rdg>
    <rdg wit="#C"/>
</app>
<app>
  <lem wit="#B"><w>the</w></rdg>
  <rdg wit="#A #C"/>
</app>
<app>
  <rdg wit="#B #C"><w>Armenian</w></rdg>
  <lem wit="#A"/>
</app>
<app>
  <lem wit="#B"><w>year</w></rdg>
  <rdg wit="#C"><w>era</w></rdg>
  <rdg wit="#A"/>
</app>
<w>401</w>
<app>
  <lem wit="#A"><w>of</w></rdg>
  <rdg wit="#B #C"/>
</app>
```

յ, Լե Կապուլան Երեւոյն Գեղեցիկ մարդկան: Որոճալով որոճեն զդիւական երե
Հայաստանէն. Ժ, Ի վեր Կապուլան Երեւոյն Գեղեցիկ մարդկան: Որոճալով որոճեն զդիւական երե

Tara L Andrews | University of Oxford | 18 November 2009

Երեւոյն Գեղեցիկ մարդկան: Որոճալով որոճեն զդիւական երե
Հայաստանէն. Ժ, Ի վեր Կապուլան Երեւոյն Գեղեցիկ մարդկան: Որոճալով որոճեն զդիւական երե

...to formatted text and apparatus

1 In the year 401 of the Armenian era...

1 In] C om.; 1 the] B the Armenian, C Armenian, A om.; 1 year] C era, A om.;
...

Գ, ԼԵ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ
ՀԱՅԱՍՏԱՆԻՆԵՅԵ. Ժ, ԻՎԵՐ ԿԱՆԿԱՆ ԵՆԵ ԴԱՆԴԵՐ ԵՐԹԱՆ ԳԵՎԵԱՏԻԿ ԺԱՐԴԿԱՆ: ՈՐՈՇԱԼՈՎ ՈՐՈՇԵՆ ՂԴԻՆԱԿԱՆ ԵՐԵ

Tara L Andrews | University of Oxford | 18 November 2009

«Ինչպէս որ, զանազան պատճառներով զկարգ ժամատեղացն: լինին և բազումք ուսում

Ռեպրեզենտացիանայ. ոչ արեղայ և ոչ կրանաւոր, ոչ իշխան և ոչ ժողովուրդ. Ի



What next?

յ, Լե Կոմպոզիցիոն Գրքան գովեստից մարդկան: Որոճալով որոճեն զդիւական երե
Հայաստանէնց. Ժ, Ի վեր հաշ

Ինչպէս Երկրորդ, Գն առաջադրուած պատորեն զկարգ ժամատեղացն: լինին և բազումք ուսում